Emotion Detection in Textual Information by Semantic Role Labeling and Web Mining Techniques

Cheng-Yu Lu[1,2], Jen-Shin Hong[1], Samuel Cruz-Lara[2]

[1] Depart of Computer Science and Information Engineering,
National ChiNan University
No.1, University Rd., Puli, Nantou, Taiwan
{cylu, jshong}@ncnu.edu.tw

[2] LORIA (UMR 7503) – University of Nancy 2
CNRS – INRIA – Universities of Nancy
Campus Scientifique – BP 239
54506 Vandoeuvre-lès-Nancy, France
{Cheng-Yu.Lu, Samuel.Cruz-Lara}@loria.fr

Abstract. Automatic emotion detection in textual information is critical for the development of intelligent interfaces in many interactive multimedia applications. In the literature, existing approaches based on keyword spotting or statistic natural language process techniques, have limited success rate in free text emotion sensing applications. In this paper, we describe a system, developed in the framework of the National ChiNan University and LORIA collaboration, that associates semantic labeling and web mining techniques, to detect several basic emotions. A common sense knowledgebase – ConceptNet – is also used in order to retrieve some additional contextual information that can be used to retrieve appropriate background images for the presentation. Our objective is to adapt a multimedia presentation by detecting emotions contained in the textual information.

1. Introduction

In the current web-based environment, to make an "affective" and aesthetic hypermedia presentation needs serious collaborating work among content experts, computer engineers and multimedia designers. This process usually is very time consuming and labor intensive.

During the past 3 years, the Digital Archive team of the National ChiNan University has devoted much effort on developing intelligent styling techniques [1] that can automatically compose Flash-based hypermedia presentations based on the emotion desired by users.

At LORIA, MLIF "Multi Lingual Information Framework" [2, 3] is being designed with the objective of providing a common conceptual model and a platform allowing interoperability among several translation and localization standards, such as XLIFF [4], TMX [5] and i18n [6], and by extension, their committed tools. MLIF is based on a methodology of standardization resulting from the ISO (sub-committees

TC37 / SC3 "Computer Applications for Terminology" and SC4 "Language Resources Management"). The asset of MLIF is the interoperability which allows experts to gather, under the same conceptual unit, various tools and representations related to multilingual data[1].

Of all the multimedia modalities, text is particularly important for sensing emotion because the majority of today's computer user interfaces are textual-based. To detect emotion out of textual data is not easy. Some research work using "Keyword-Spotting" [7, 8], "Lexical affinity" techniques to extract the emotion [9], but the results are not satisfactory. In this article, we will present a novel approach of emotion detection using based on shallow semantic analysis of English textual content. A common sense knowledgebase – ConceptNet – is also used in order to retrieve some additional contextual information that can be used to retrieve appropriate background images for the presentation. In the following section, the underlying approach and an example application of an affective chatting room will be described.

As a collaborating work between the National ChiNan University and LORIA, the next step is to develop techniques that automatically sense the underlying emotion from multimedia and multilingual contents to invoke corresponding hypermedia presentations with correct styles. In particular, we would like to:

1. Enhance the current mechanism to detect emotion out of textual data;
2. Apply the MLIF as a bridge allowing to process multilingual textual data;
3. Study the mechanisms for intelligent styling system to make presentations with styles corresponding to the correct emotions.

As a first step of our collaboration, we have developed a prototype system that associates a semantic labeling tool and a web mining engine, in order to detect emotions (i.e. happiness, sadness, anger, fear, surprise, …). It should be noted that, currently, multilingual and intelligent styling aspects have not yet been implemented in the proposed prototype system. These important aspects of our research work will be taken progressively into account.

2. Literature review.

In the literature, there are a number of approaches to textual emotion detection. Keyword spotting approaches apparently can't apply to sentences without clear affective keywords. Lexical affinity techniques classify the emotion of a linguistic unit based on the affinity of the linguistic unit and an affective keyword. For example, if a phrase appears closely much more often with the work "happy" than "sad", it is reasonable to believe that the emotion associated this phrase is happy. Such approach requires a

[1] It should also be noted that, within the framework of ITEA "Passepartout" project [10], we are experimenting with some basic scenarios where MLIF is associated to XMT (eXtended MPEG-4 Textual format [11]) and to SMIL (Synchronized Multimedia Integration Language [12]). Our main objective in this project is to associate MLIF to multimedia standards (e.g. MPEG-4, MPEG-7, and SMIL) in order to be able, within multimedia systems, to represent and to handle multilingual content in an efficient, rigorous and interactive manner.

really large scale corpus and a really sophisticated search engine to get statistically significant results. The search engine need not only retrieves all documents with the phrases but also need to determine the distance of the two phrases. Without the distance information, the approach can be trapped easily. For example, a query to Google using "war, happy" gets many more pages than "war, sad".

Alm [13] proposed a sentence level emotion classifier by feeding a machine learning algorithm in a training corpus of affectively annotated texts. The approach take into account a number of features such as the emotion words, punctuation, and story progress, thematic story type, etc. The primitive results appear to be promising when given a sufficiently large text input within certain narrative structures. However, the applicability of such emotion sensing to domain-independent text in sentence level has to be investigated further. Hugo [14] provides a novel approach to detect sentence-level emotion based on a large-scale common sense knowledgebase. The approach inherits affective nature of everyday situations (such as "getting into a car accident") to classify sentences into emotion categories. Essentially, a number of affect keywords act as the "emotion grounds" for assigning the sentences with clear affect keywords. Emotion of sentences without clear affect keywords are assigned automatically by propagating the emotion with a certain relationship between. The accuracy of such propagation process has not yet been investigated. Also, limited coverage of the knowledge based present limitation of such approach in more general applications.

Emotion theories, particularly that of Ortony, Clore and Collins [15], have been used widely in order to detect emotion within interaction systems [16, 17, 18]. Through a variety of hand-crafted models of affect, based on psychological theories, such approach specifies how events, agents and objects are appraised according to an individual's goals, standards and attitudes. Because of the thorough nature of the approach, its application in free text requires a deep understanding and analysis of the text which require deep semantic parsing technology that is rather beyond the reach of current naturally language processing technologies. On the other hand, over the past few years, the advancement of the shallow semantic parsing tools actually reaches to an acceptable accuracy rate in English sentences [19]. Before the birth of a really sophisticated deep semantic parser, we argue that the emotion theories should be modified to some extent to fully take advantage of the state-of-art semantic labeling techniques for emotion detection. In the following section, we will explain the general idea and primitive results of our methodology. Note that, each of the above mentioned existed approaches might find their best applications in certain scenarios. A real through system should combine those existed approaches to achieve a higher accuracy rate.

3. A System for Emotion Detection

Currently, we are building a system that uses a semantic role labeling tool in order to detect emotions within textual information. In order to validate our ideas, we are only working with English. Later, by using MLIF, the system should also include French and Chinese.

In our current emotion detection system, we are using two publicly available tools: a semantic labeling tool developed by the Cognitive Computation Group of the University of Illinois at Urbana-Champaign [20], and a web mining engine as Google [21].
The semantic labeling tool implements a semantic role labeling (SRL) in which, for each verb in a sentence, the goal is to identify all constituents that fill a semantic role, and to determine their roles, such as Agent, Patient or Instrument, and their adjuncts, such as Locative, Temporal or Manner.
By using this tool, some input text may be parsed and labeled with several different components: "subjects", "objects", "verbs", etc. Figure 1(a) shows the labeling of the input text: "A girl met a tiger". The semantic role labeling output gives a three parts result:

1. "A girl" is tagged with A0 (i.e. subject);
2. "met" is tagged with V:met ;
3. "a tiger" is tagged with A1 (i.e. object).

Figure 1(b) shows more some tags and their corresponding annotations in the semantic labeling tool.

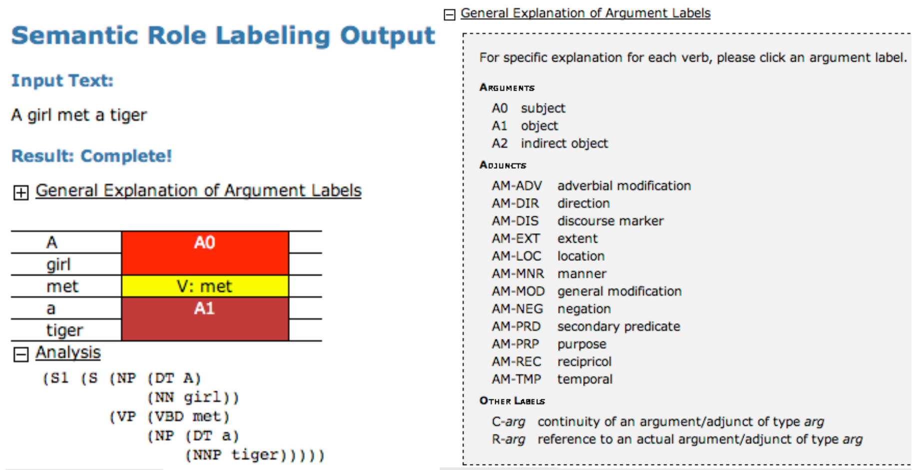

Figure 1. "Semantic Role Labeling of the sentence "A girl met a tiger" ".

A web mining engine (i.e. "Google") is also used within the proposed system. A web mining engine allows searching from a specific keyword and provides purely empirical answers to lexical questions. Google's "define" function is a service which gives up-to-date definitions automatically gathered from the internet. Figure 2 and Figure 3 shows the definitions of "Tiger" and "Wolf" respectively.

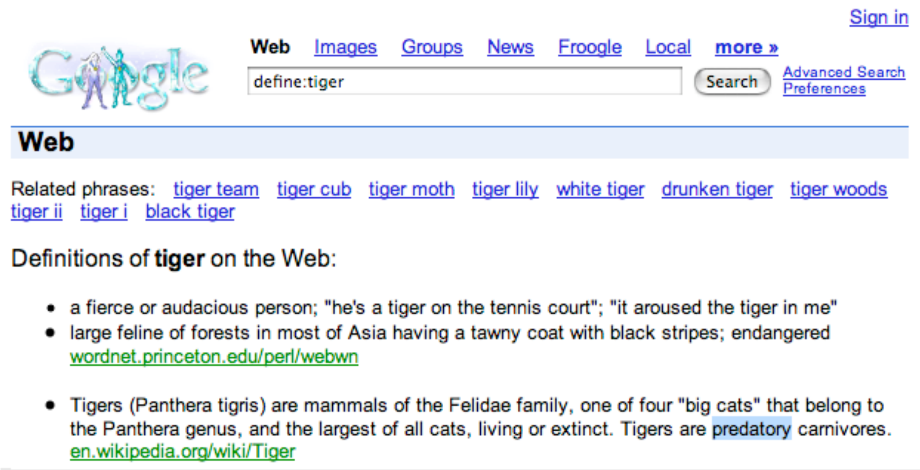

Figure 2. “Google’s definition of “Tiger””.

One of the definitions of “Tiger” mentions that: “Tigers are predatory carnivores”. We also found that, one of the definitions of “Wolf” indicates “any of various predatory carnivorous canine mammals of North America…”.

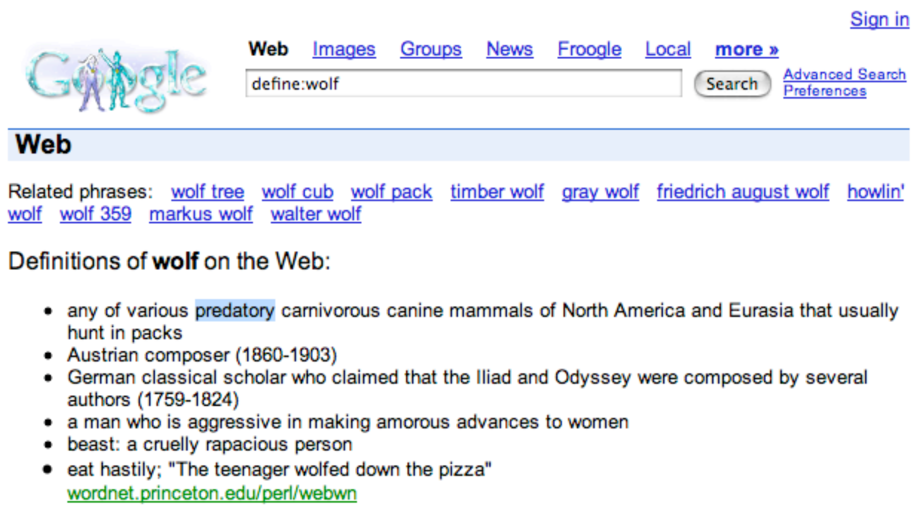

Figure 3. “Google’s definition of “Wolf””.

From the two definitions above, we find that the adjective “predatory” appears in the definitions of “Tiger” and “Wolf”. Next, we save the specific adjective “predatory” into a table with some code like “ADJ_#”. If the system detects different words containing the same definition, it will reference them by using the same adjective. So far, we have only considered 15 different adjective categories (see Table 1). We will increase the number of adjective categories progressively.

id	adjective	adj code	id	adjective	adj code
29	illicit	ADJ_13	6	youthful	ADJ_3
30	illegitimate	ADJ_13	7	carnivorous	ADJ_5
31	criminal	ADJ_13	4	predatory	ADJ_5
32	felonious	ADJ_13	5	rapacious	ADJ_5
25	unlawful	ADJ_13	13	reptile	ADJ_7
37	difficult	ADJ_14	17	hazardous	ADJ_8
38	hard	ADJ_14	8	lacking	ADJ_9
39	rough	ADJ_14	83	precious	ADJ_20
40	laborious	ADJ_14	33	awful	ADJ_18

Table 1. “Some adjective categories and their corresponding code numbers”.

Our approach differs from other related research work significantly. Some methods can only perform classification to tell which emotional category a sentence belongs to by detecting the emotional keywords out of the sentences [7, 8]. Our method takes into account some semantic issues that help to analyze sentences like “A girl met a tiger”. We are currently working only with seven different emotions: “Happiness”, “Sadness”, “Anger”, “Fear”, “Disgust”, “Surprise” and “Neutral”.
So, how do we detect emotions? Let’s go back to the sentence “A girl met a tiger”. When we analyze (i.e. Semantic Labeling Role) this sentence, we obtain that: “A girl” is the <u>subject</u> of the sentence, “met” is the <u>verb</u> of the sentence, and “a tiger” is the <u>object</u> of the sentence. Next, subjects and objects are linked to adjectives respectively based on their Google-define results. In the sentence “A girl met a tiger”, the subject (i.e. “A girl”) is linked to the adjective “youthful” (i.e. ADJ_3), and the object (i.e. “a tiger”) is linked to the adjective “predatory” (i.e. ADJ_5). Finally, we combine ADJ_3 and ADJ_5 by using the verb “meet”. The result is a “Fear” emotion (see Figure 4 and Table 2).

In our current prototype system, the primitive emotion model is created using a manual procedure. Table 2 shows a partial list of the emotion rules for different combination of some selected adjectives and a verb. If we take the record (id:35) as an example (i.e. Girl buy jewel), the subject (i.e. Girl) and the object (i.e. jewel) of this sentence are related to adjectives “ADJ_3” (i.e. youthful) and “ADJ_20” (i.e. precious) respectively. If we combine these adjectives with the verb “buy” (or “purchase”, “shop”), we obtain a “happy” emotion.

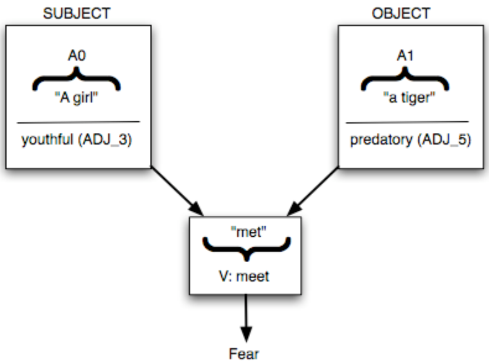

Figure 4. “Combining subjects and objects with verbs for emotion detection”.

id	verb	combination	emotion	Example
35	buy	ADJ_3:ADJ_20	happy	Girl buy jewel
22	abandon	ADJ_3:ADJ_8	happy	Girl abandon smoking
23	kill	ADJ_5:ADJ_3	sad	Tiger kill girl
25	kill	ADJ_3:ADJ_5	surprised	Girl kill tiger
34	compete	ADJ_3:ADJ_5	surprised	Girl compete tiger
36	ease	ADJ_3:ADJ_8	happy	Girl ease smoking
30	hate	ADJ_3:ADJ_16	angry	Girl hate tsunami
31	hate	ADJ_3:ADJ_5	angry	Girl hate tiger
4	meet	ADJ_3:ADJ_7	fear	Girl meet tiger
6	meet	ADJ_5:ADJ_3	fear	Tiger meet girl
7	meet	ADJ_9:ADJ_3	sad	Poor meet girl
8	meet	ADJ_3:ADJ_9	sad	Girl meet poor
9	meet	ADJ_3:ADJ_5	fear	Girl meet tiger
33	compete	ADJ_3:ADJ_16	surprised	Girl compete tsunami.

Table 2. “Combining adjectives and verbs”.

Also, in order to be able to handle more verbs, a table which contains verbs and their synonyms has been constructed (see Table 3). By establishing this table, the system will handle more verbs which have the same attribute with the same rules. For example, we consider that “abandon”, “leave”, and “cease” are synonyms. So far, we have

13 classifications in our system. The synonyms in Table 3 are referred from the on-line Yahoo-Dictionary (http://tw.dictionary.yahoo.com/).

id	verb	group	id	verb	group
5	discard	abandon	28	hate	hate
12	acquire	acquire	29	dislike	hate
13	gain	acquire	30	despise	hate
14	earn	acquire	1	destroy	kill
15	secure	acquire	8	kill	kill
16	obtain	acquire	9	slay	kill
24	astonish	astonish	10	slaughter	kill
25	surprise	astonish	11	murder	kill
26	amaze	astonish	6	meet	meet
27	astound	astonish	50	encounter	meet
31	buy	buy	51	join	meet
32	shop	buy	7	see	meet

Table 3. "Some verbs and their synonyms".

So, the Semantic Role Labeling of the sentence "A girl met a tiger" allows to detect a "Fear" emotion.

In our prototype system, Concept Net is just used in order to retrieve some additional information (i.e. Location, for example). Actually, we can use Concept Net in order to find a place associated to "tiger". Figure 5 shows a request sent to Concept Net.

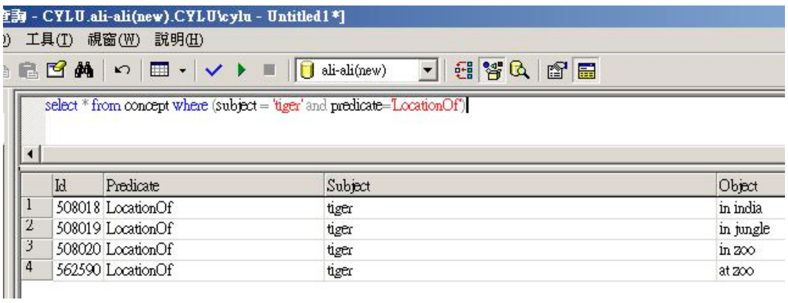

Figure 5. Sending a request to Concept Net.

So, one of the possible locations (i.e. LocationOf) related to "tiger" is "in jungle". This information may be useful, essentially, in order to change the background image or the background audio of a multimedia application.

4. A chatting room application.

In order to illustrate our proposals, we have built an “affective chatting room”. Figure 6 shows the general architecture of this application.

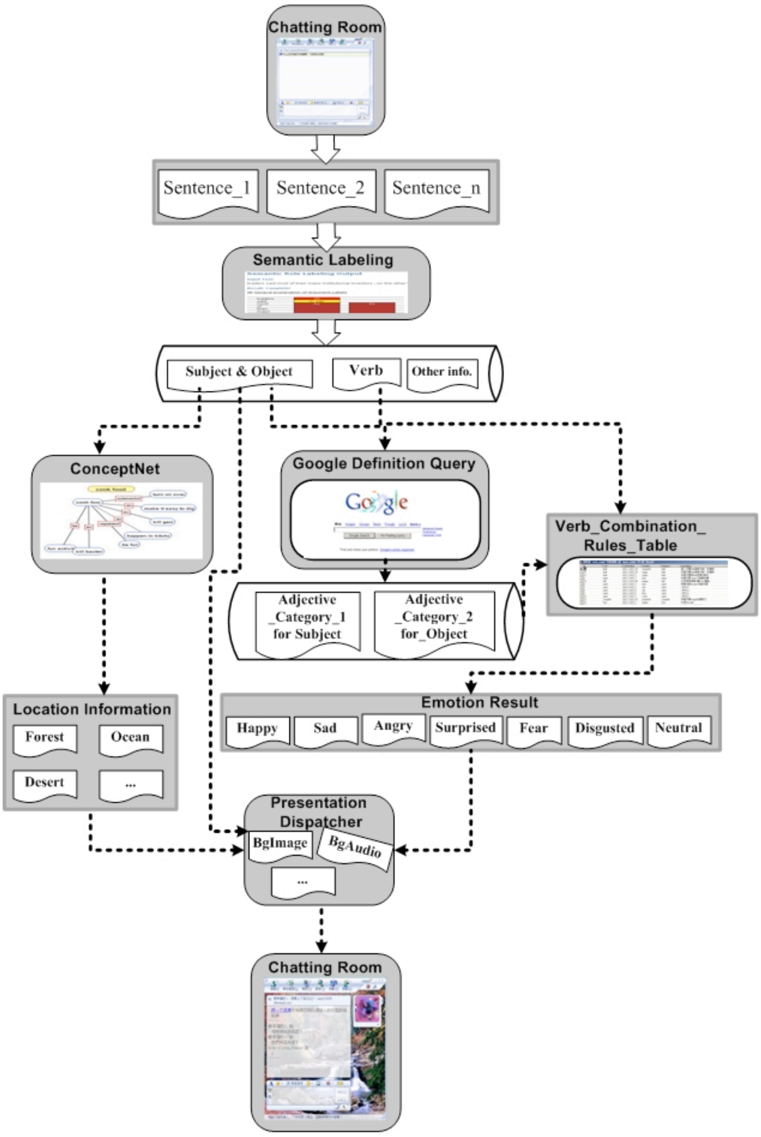

Figure 6. The general architecture of a “chatting room application”.

This kind of application has been frequently used in order to illustrate emotion detection from textual information [22, 23]. However, it should be noted that, even if the chatting room application we have developed is similar to those of [22, 23], we have not used the same models, techniques and methods. In these systems [22, 23] emotions are detected by means of “Keyword-Spotting”. In our system, emotions are detected by combining subjects and objects with verbs. Subjects and objects are identified by associating results obtained from: a semantic labeling tool and a Google’s define operation. Finally, ConceptNet requests are used in order to retrieve some additional information (i.e. Location, for example).

Figure 7 shows the user interface of our chatting room application. On the right-hand side of this interface one should note that the analysis (i.e. Semantic Role Labeling) of the sentence “It’s a good car” leads to the “happiness” emotion. Detecting this emotion may allow to change, for example, the background image and color of the user interface.

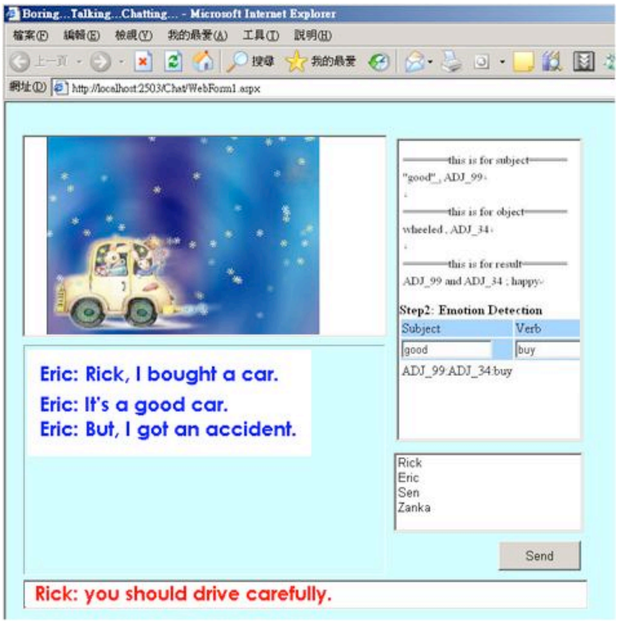

Figure 7. The chatting room application.

To present the sentences in the chatting room, we use an intelligent styling engine developed by ChiNan university to properly adapt the hypermedia presentation based on the emotion of the sentence. In the overall framework, intelligent styling engine will receive the information “subject” and “object” from the semantic labeling procedure and emotion results of each sentence. A corresponding presentation containing background images, background audios, and others with appropriate style will be selected.

5. Conclusion and future work.

In the framework of the National ChiNan University and LORIA collaboration, we have presented a prototype system that associates a semantic labeling tool to a web mining engine in order to detect emotions. A common sense knowledgebase as, Concept Net, may also be used in order to retrieve some additional information that can be used to retrieve appropriate background images for the presentations.

To make the emotion detection system more robust, in the future there are a number of issues we intend to explore further.

1. First, more argument labels within the semantic labeling techniques should be taken into account in the emotion model;
2. Second, more adjective groups need to be included in the emotion model to account for more emotions.
3. Third, it should be noted that, frequently, in order to obtain an accurate emotion for a sentence, the context in which a sentence is used must be taken into account. We intend to incorporate various "story-understanding" techniques to take account of the context information for emotion detection;
4. Finally, we intend to apply to other languages, such as French, Chinese, Spanish, … etc. with the assistance of MLIF.

6. Acknowledgement

This work is supported by the National Science Council of Taiwan, under grant NSC-94-2422-H-260-002, and the doctoral student program numbered NSC- 094-JFA-12-001. The authors wish to thank Professor Dan Roth of the Department of Computer Science of University of Illinois at Urbana-Champaign for allowing us to access the semantic labeling tools; and many graduate students at the Digital Archive Lab of ChiNan University for contributing to the development of the affective hypermedia chatting room.

7. References.

1. Jen-Shin Hong, Bai-Hsuan Chen, Sheng-Hao Hung, Jieh Hsiang. Towards an Integrated Digital Museum System- the ChiNan Experiences. Journal of Digital Libraries (Springer). Special Issue on Digital Museum, in Print.
2. S. Cruz-Lara, S. Gupta, L. Romary. Handling Multilingual Content in Digital Media: The Multilingual Information Framework. EWIMT, 2004 European workshop on the Integration of Knowledge, Semantics and Digital Media Technology.
3. S. Cruz-Lara, S. Gupta, J.D. Fernández García, and L. Romary. Multilingual Information Framework for Handling Textual Data in Digital Media. IEEE AMT 2005, The Third International Conference on Active Media Technology. Takamatsu, Kagawa, Japan. 19-21 May, 2005.

4. XLIFF. Oasis (2003). XML Localisation Interchange File Format, http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=xliff. Accessed 2003-10-31
5. TMX. Oscar / Lisa (2000) Translation Memory eXchange, http://www.lisa.org/tmx. Accessed 2001-06-27
6. i18n. W3C (2003) Internationalization (i18n). http://www.w3.org/International/.
7. Elliott, C. The Affective Reasoner: A Process Model of Emotions in a Multi-agent System. PhD thesis, Northwestern University, May 1992. The Institute for the Learning Sciences, Technical Report No. 32.
8. Zhe, X. & Boucouvalas, A. C. 2002. Text-to-Emotion Engine for Real Time Internet Communication. In Proceedings of International Symposium on Communication Systems, Networks and DSPs, Staffordshire University, UK, pp 164-168.
9. Jonathon Read, Recognising Affect in Text using. Pointwise-Mutual Information. Master of Science. University of Sussex, UK. 2004.
10. ITEA "Information Technology for European Advancement". Passepartout project "Exploitation of advanced AV content protocols (MPEG 4/7)" ITEA 04017. http://www.itea2.org/public/project_leaflets/PASSEPARTOUT_profile_oct-05.pdf
11. XMT. extended MPEG-4 Textual format. ISO/IEC FCD 14496-11, Information technology -- Coding of audio-visual objects -- Part 11: Scene description and application engine; ISO/IEC 14496-11/Amd 4, XMT & MPEG-J extensions.
12. Synchronized Multimedia Integration Language (SMIL 2.0) . World Wide Web Consortium. http://www.w3.org/TR/smil20/
13. C. O. Alm and D. Roth and R. Sproat, Emotions from text: machine learning for text-based emotion prediction, Proceedings of HLT/EMNLP, 2005
14. Hugo Liu, Henry Lieberman, Ted Selker. (2003). A Model of Textual Affect Sensing using Real-World Knowledge. Proceedings of the 2003 International Conference on Intelligent User Interfaces, IUI 2003, 125-132. ACM Press.
15. Ortony A., Clore, G. L., and Collins, A. The cognitive structure of emotions. New York: Cambridge University Press. 1988.
16. Dyer, M.G. Emotions and Their Computations: Three Computer Models. Cognition and Emotion, Vol. 1, No. 3, 323-347, 1987.
17. Prendinger, H. & Ishizuka, M. 2001. Simulating Affective Communication with Animated Agents. In Proceedings of Eighth IFIP TC.13 Conference on Human-Computer Interaction, Tokyo, Japan, 182-189.
18. Bartneck, C. (2002). Integrating the OCC Model of Emotions in Embodied Characters. Proceedings of the Workshop on Workshop on Virtual Conversational Characters: Applications, Methods, and Research Challenges.
19. V. Punyakanok, D. Roth, W. Yih, and D. Zimak, Semantic Role Labeling via Integer Linear Programming Inference. Proc. the International Conference on Computational Linguistics (COLING) (2004) pp. 1346—1352.
20. Cognitive Computation Group. University of Illinois at UIrbana-Champaign. Software & Demos. Semantic Role Labeling Demo. http://l2r.cs.uiuc.edu/~cogcomp/srl-demo.php
21. Google. http://www.google.com
22. JunHee Park. Mood Chat Emotional Interface. Thesis documentation. Parsons School of Design - Master of Fine Arts in Design and Technology. May 2001.
23. Chunling Ma, Helmut Prendinger, Mitsuru Ishizuka. A Chat System Based on Emotion Estimation from Text and Embodied Conversational Messengers. ICEC 2005, The Seventh International Conference on Electronic Commerce. Xi'an, China. August 2005.